FTML Exercices 4

TABLE DES MATIÈRES

1 CONVEXITÉ

— Connaître les définitions de la section 2.3.1 de lecture_notes.pdf et la proposition 8 de la section 2.3.8.

— Essayer de prouver les exemples de propriétés indiquées dans le paragraphe **Examples**.

Pour certains, nous avons déjà vu les solutions dans les exercices 2. Pour certaines propriétés, c'est plus simple en considérant les dérivées ou les gradients (proposition 8 de la section 2.3.2)

2 REGRESSION LOGISTIQUE

We consider the logistic regression problem, in the following setting :

— $\mathcal{X} = \mathbb{R}^d$

— $\mathcal{Y} = \{-1, 1\}$

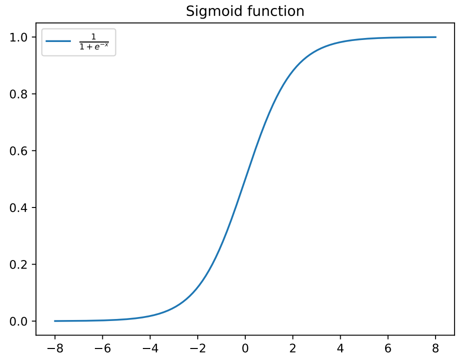
— Logistic loss :

$$l(\hat{y}, y) = \log(1 + e^{-\hat{y}y}) \tag{1}$$

We also define the following function :

Definition 1. Sigmoid function
$\sigma : \mathbb{R} \to \mathbb{R}$.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$



1] Show that σ is differentiable and that

$$\forall z, \sigma'(z) = \sigma(z)\sigma(-z) \tag{3}$$

2] Show that $l(\hat{y}, y)$ is strictly convex in its first argument, which means for fixed y, $\hat{y} \mapsto l(\hat{y}, y)$ is strictly convex. Using properties that relate convexity and the derivative of a function, and using the sigmoid function will be helpful.

3] Without regularization, the empirical risk writes :

$$R_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} l(x_i^T\theta, y_i) \tag{4}$$

a Show that $\theta \to R_n(\theta)$ is convex

b Compute the gradient $\nabla_\theta R_n(\theta)$ of the empirical risk $R_n(\theta)$ in this setting.